

**Brain networks underlying the differences in audiovisual integration for reading
between children and adults and its disruption in dyslexia**

Junjun Li^{1, 2†}, Yang Yang^{1, 2†}, Nestor Viñas-Guasch³, Yinghui Yang^{1, 2}, Hong-Yan Bi^{1, 2*}

¹ CAS Key Laboratory of Behavioral Science, Center for Brain Science and Learning Difficulties, Institute of Psychology, Chinese Academy of Sciences, Beijing, 100101, China

² Department of Psychology, University of Chinese Academy of Sciences, Beijing, 100049, China

³ Clinical Brain Lab, Nanyang Technological University, Singapore, Singapore

[†]These authors have contributed equally to this work.

*Please send correspondence to:

Hong-Yan Bi, Ph.D.

Institute of Psychology, Chinese Academy of Sciences

Beijing 100101, China

Email: bihy@psych.ac.cn

Abstract

Building robust letter-to-sound correspondences is a prerequisite for reading, and such audiovisual integration becomes progressively automatic with development. However, the neural mechanisms underlying the development of audiovisual integration for reading are largely unknown. This study used functional magnetic resonance imaging (fMRI) in a lexical decision task to investigate the changes of brain functional networks that support audiovisual integration for reading between normally developing children (9-12 years old) and adults (20-28 years old). The identified networks were further examined in children with developmental dyslexia (9-12 years old). Results revealed that adults enhanced connectivity in a prefrontal-superior temporal network relative to children, reflecting the attentional modulation to the development of audiovisual integration. Moreover, this network was disrupted in dyslexics, confirming its essential role in audiovisual integration for reading. This study, for the first time, elucidates the neural basis underlying the development of audiovisual integration for reading.

Key words audiovisual integration; reading; development; brain network; fMRI

1 Introduction

2 Establishing reliable and robust associations between visual and auditory information
3 is the foundation of reading acquisition and development (Blau et al., 2010; Blau, van
4 Atteveldt, Ekkebus, Goebel, & Blomert, 2009; Holloway, van Atteveldt, Blomert, &
5 Ansari, 2015; van Atteveldt, Formisano, Goebel, & Blomert, 2004). Dysfunction in
6 integrating orthographic and phonological information into a unified audiovisual
7 percept has been identified as a vital factor of reading failure in alphabetic languages
8 (Blau et al., 2010; van Atteveldt et al., 2004) and logographic languages (Yang, Yang,
9 Li, Xu, & Bi, 2020). Congruent phonological information and visual scripts
10 complement each other, thus improving the accuracy and responsiveness of visual
11 word recognition (Kast, Bezzola, Jäncke, & Meyer, 2011; Raij, Uutela, & Hari, 2000).

12
13 At the neural level, numerous neuroimaging studies using different task paradigms
14 have identified several brain regions involved in audiovisual integration for reading.
15 For example, by comparing brain activation between the audiovisual responses and
16 the summations of unisensory responses, Raij et al. (2000) and van Atteveldt et al.
17 (2004) found that the bilateral superior temporal gyrus/superior temporal sulcus
18 (STG/STS), the left frontoparietal region and the right frontal cortex were engaged in
19 the integration of letters and speech sounds in skilled adult readers (Raij et al., 2000;
20 van Atteveldt et al., 2004). In addition, several studies have examined the brain
21 substrates of audiovisual integration by the congruency effect, which is based on the
22 comparison of brain responses between congruent and incongruent audiovisual

1 stimulus pairs (Blau, van Atteveldt, Formisano, Goebel, & Blomert, 2008; W. Xu,
2 Kolozsvári, Oostenveld, Leppänen, & Hämäläinen, 2019). The studies of adult
3 readers found that the bilateral STG/STS, Heschl sulcus/planum temporale,
4 middle/inferior temporal gyrus (MTG/ITG), middle/inferior frontal gyrus (MFG/IFG),
5 cingulate gyrus (CG), superior parietal lobule (SPL) and fusiform gyrus (FuG) were
6 engaged in audiovisual integration processing (Blau et al., 2008; Holloway et al.,
7 2015; van Atteveldt, Blau, Blomert, & Goebel, 2010; van Atteveldt, Formisano,
8 Blomert, & Goebel, 2006; W. Xu et al., 2019). In short, brain activation in the
9 superior temporal cortex and frontal cortex were consistently observed to support the
10 audiovisual integration in reading processing.

11
12 With practice and reading development, audiovisual integration processing becomes
13 gradually automatic and optimal (Beierholm & Adams, 2016; Froyen, Bonte, van
14 Atteveldt, & Blomert, 2009). A study of Dutch children showed that the processing
15 time for letter-speech sound associations steadily decreased over the full range of
16 primary school grades, despite early acquisition of associations between orthography
17 and phonology (Blomert & Vaessen, 2009), suggesting an ongoing development
18 towards automatic processing. Furthermore, Froyen et al. (2009) found that skilled
19 adult readers, but not children (10-12 years), showed enhanced mismatch negativity
20 (MMN) amplitude of letter-speech sound presented simultaneously than presented
21 separately (Froyen et al., 2009), suggesting the differences in neural activity of
22 audiovisual integration in reading between children and adults. To the best of our

1 knowledge, only an early study examined the differences in effective connectivity of
2 audiovisual integration between children and adults (Dick, Solodkin, & Small, 2010),
3 but that study examined the audiovisual integration in the context of speech
4 comprehension. Consequently, it is unknown about the differences in the brain
5 activity of audiovisual integration for visual word recognition (reading) between
6 children and adults.

7

8 Recent evidence shows that audiovisual integration requires interplay between
9 distributed regions (Calvert, 2001; Driver & Noesselt, 2008; Paraskevopoulos,
10 Kraneburg, Herholz, Bamidis, & Pantev, 2015). Specifically, audiovisual integration
11 recruits high-level cognitive processes (e.g. attention and semantic processing)
12 resulting in its late development (Barutcu et al., 2010; McNorgan, Randazzo-Wagner,
13 & Booth, 2013; Murray & Wallace, 2012; Talsma, Senkowski, Soto-Faraco, &
14 Woldorff, 2010), and thus regions involved in higher order processing might interact
15 with regions involved in audiovisual integration. In this context, a large-scale
16 functional network analysis may be a more informative method to understand the
17 brain organization underlying audiovisual integration in reading and its development
18 from childhood to adulthood. Functional networks are typically modeled as graphs
19 composed of nodes (the cortical regions contributing to a network) and edges (the
20 connections between nodes) (Meunier, Achard, Morcom, & Bullmore, 2009;
21 Paraskevopoulos et al., 2015). Previous studies have successfully applied network
22 analysis method to unveil the neurodevelopment of functional networks for reading

(X. Liu et al., 2018) and for expressive language ability (Doesburg, Tingling, MacDonald, & Pang, 2016).

Using functional network analysis, the present study aimed to unveil the changes in neural mechanisms underlying the audiovisual integration for reading between children and adults. First, we compared the brain networks of audiovisual integration between normally developing child readers (9-12 years old) and skilled adult readers (20-28 years old). Following previous studies (Kast et al., 2011; Yang et al., 2020), a lexical decision task was used to examine audiovisual integration in a real reading context. Participants were asked to decide whether the visual symbols presented simultaneously with congruent or incongruent speech sounds were real Chinese characters. The congruency effect was adopted as the index of audiovisual integration effect (Blau et al., 2010; Blau et al., 2009). Our hypothesis was that compared to children, skilled adult readers would show greater functional connectivity in a widespread network involving the core regions of audiovisual integration (such as STG, MFG and IFG) and the higher-order association cortices (such as prefrontal and parietal cortices).

Afterwards, the identified functional networks that differed between the two age groups were explored in a sample of children with developmental dyslexia. The rationale was that if the abovementioned functional networks are critical to audiovisual integration development for reading, we would expect to observe the

disruption of these functional networks in individuals with dyslexia.

Materials and Methods

Participants

Twenty-five normally developing children (9 females, mean age = 11.45 ± 0.83 years, abbreviated as CH), twenty-one adults (11 females, mean age = 23.85 ± 2.61 years, abbreviated as AD) and fourteen children with dyslexia (4 females, mean age = 10.99 ± 1.03 years, abbreviated as DD) participated in this study. All participants were native Mandarin Chinese speakers, and were right-handed assessed by the Handedness Inventory (Department of Neurology, Beijing Medical University Hospital). All participants had normal hearing, normal or corrected-to-normal vision and were not suffered from ophthalmological or neurological abnormalities.

The sample size of CH and AD was determined a priori using G*Power (Version 3.1, <http://www.gpower.hhu.de/>) (Faul, Erdfelder, Lang, & Buchner, 2007), which indicated that a total of 34 participants were required for a medium partial η^2 of 0.06 (effect size $f = 0.25$) and a power of 0.8 with an alpha of 0.05 (Mumford, 2012; Simonet, Roten, Spierer, & Barral, 2019). In order to compensate for potential exclusion of participants (e.g. excessive head motion), we recruited more than 20 participants for each group.

The dyslexic participants were from one published study (Yang et al., 2020). The

screening criteria includes: 1) having a reading score at least one and a half standard deviations below the average score of children in the same grade assessed by the Character Recognition Measures and Assessment Scale (CRM) (X. L. Wang & Tao, 1996); 2) having normal score (above 85) of the standard score of non-verbal intelligence quotient (IQ) evaluated by Combined Raven's Progressive Matrices (CRT) (Li, Chen, & Jin, 1989); and 3) not suffering from ADHD assessed by the Chinese Classification of Mental Disorder 3 (CCMD-3).

The study was approved by the ethics committee of the Institute of Psychology, Chinese Academy of Sciences. Each adult participant and child participant's guardian gave written informed consents prior to this study. Demographic information and the results of screening tests are shown in Table 1.

Linguistic-cognitive tests

All participants were administered three linguistic-cognitive tests measuring reading accuracy, reading fluency and phonological awareness, respectively. The reading accuracy test consists of 172 Chinese characters with varying word frequency. Participants were required to name overtly all the characters as accurately as possible, with no time limit. The reading fluency test consisted of 160 high and medium frequency Chinese characters. Participants were asked to read these characters aloud as quickly and accurately as possible within one minute. In both tests, one point was awarded for each character that was read correctly. In the phonological awareness test,

1 participants heard three syllables, one of which was different from the others in
 2 consonant, vowel or tone (10 items for each type). Participants were asked to select
 3 the syllable that differed from the others. One point was given for each correct
 4 judgment. Three CH, six AD and two DD did not participate in the reading tests, so
 5 their scores on the linguistic-cognitive tests were missing. The test scores of the
 6 remaining participants are presented in Table 1.

7
 8 **Table 1** Demographic information and performance in the reading tests for all three
 9 groups of participants

	CH (n=25)	AD (n=21)	DD (n=14)	<i>p</i> -values (t-tests or χ^2 -tests)	
				CH vs. AD	CH vs. DD
Age	11.45 (0.83)	23.85 (2.61)	10.99 (1.03)	< 0.001	0.139
Male/Female	16/9	10/11	10/4	0.264	0.637
Raven IQ	119.68 (15.72)	–	106.71 (14.88)	–	0.016
Reading score	2849 (260)	–	2007 (293)	–	< 0.001
Reading accuracy	104.73 (11.37)	147.47 (8.01)	84.75 (17.63)	< 0.001	< 0.001
Reading fluency	105.23 (15.06)	141.13 (18.00)	69.08 (21.35)	< 0.001	< 0.001
Phonological awareness	24.55 (0.86)	24.80 (0.80)	15.17 (1.51)	0.838	< 0.001

10 Notes: Data are presented as the mean (standard deviation). CH = normally developing children, AD =
 11 adults and DD = children with dyslexia.

12

13 **Stimuli and task design**

The visual stimuli were 30 high-frequency Chinese characters and 30 pseudocharacters. All the real characters were compound characters composed of a phonetic radical and a semantic radical. The pseudocharacters were created by combining a phonetic and a semantic radical together in their legal positions in Chinese orthography, but were unpronounceable and nonsensical. The visual complexity (stroke number and frequency of radicals) of real characters and pseudocharacters was matched (see Supplementary Materials Table S1). The auditory stimuli consisted of 120 pronunciations of Chinese characters recorded by a female native speaker. Stimuli were assigned to five experimental conditions: the condition of audiovisually congruent characters (AVcon) in which a real character and its sound were presented simultaneously; the condition of audiovisually incongruent characters (AVincon) in which a real character and a incongruent sound were presented; the condition of audiovisually pseudocharacters (AVpseudo) in which a pseudocharacter appeared with a sound of a real character; the condition of visual real characters (Vreal) and visual pseudocharacters (Vpseudo) in which a real character/pseudocharacter was presented visually in isolation.

An event-related design was adopted for fMRI scan. Each participant underwent two runs. Each run included 15 trials of AVcon, 15 trials of AVincon, 30 trials of AVpseudo, 15 trials of Vreal, 15 trials of Vpseudo, together with 47 null trials, for a total of 137 trials, presented in pseudo-random order. In each task trial, a fixation was first presented in the center of the screen for 500 ms, followed by the presentation of

stimuli for 1200 ms and a blank screen for 800 ms. The visual stimuli presented 1200 ms and the auditory stimuli presented simultaneously with visual stimuli (in the bimodal conditions), lasting from 185 to 509 ms. Each null trial consisted of 500 ms of fixation and 2000 ms of a blank screen. Following previous studies (Kast et al., 2011; Yang et al., 2020), participants were instructed to complete a lexical decision task, in which they were required to attend to the visual stimuli and determine whether they were real Chinese characters or not by pressing buttons.

Image acquisition

All participants underwent scanning at a 3T MRI Siemens Prisma^{fit} scanner at the Beijing MRI Center for Brain Research of the Chinese Academy of Sciences. Functional MRI time series data were obtained using a BOLD-sensitive T2*-weighted gradient-echo echo planar imaging (EPI) sequence (32 slices, slice thickness = 3 mm with a 0.6-mm gap, in-plane resolution = 3 mm × 3 mm, flip angle = 90°, repetition time = 2500 ms and echo time = 30 ms). High spatial resolution anatomical images were acquired using a T1-weighted, magnetization-prepared rapid acquisition gradient echo (MPRAGE) sequence (slice thickness = 1 mm, in-plane resolution = 1.0 mm × 1.0 mm, flip angle = 8°, repetition time = 2600 ms and echo time = 3.02 ms).

fMRI data analysis and statistics

Preprocessing

Image preprocessing was conducted using SPM8 (<http://www.fil.ion.ucl.ac.uk/spm/>,

1 Wellcome Department of Cognitive Neurology, University College London, London).
2 The fMRI time series data were firstly corrected for slice timing and head motion, and
3 then normalized into Montreal Neurological Institute (MNI) stereotactic space with
4 cubic voxels at $2\text{ mm} \times 2\text{ mm} \times 2\text{ mm}$ spatial resolution. Finally, the normalized
5 functional images were smoothed with an isotropic Gaussian kernel with a 6 mm
6 full-width at half-maximum. Seven CH and two DD are excluded from the following
7 analysis because of excessive head motion during the scanning period ($> 3\text{ mm}$
8 translation or $> 3^\circ$ rotation), thus the final sample size was 18 CH, 21 AD and 12 DD
9 subjects.

10

11 **Functional network analysis**

12 **Creation of functional connectivity matrices**

13 A total of 264 functional regions of 10 mm diameter spheres were selected as nodes
14 based on an validated parcellation template (Power et al., 2011; Vatansever, Menon,
15 Manktelow, Sahakian, & Stamatakis, 2015; Wagner et al., 2019). Functional
16 connectivity (FC) matrices were created with the CONN Functional Connectivity
17 Toolbox (Whitfield-Gabrieli & Nieto-Castanon, 2012). Specifically, the blood oxygen
18 level dependent (BOLD) time series corresponding to the AVcon and AVincon
19 conditions were first extracted and concatenated separately over trials. Nuisance
20 BOLD signal fluctuations from cerebrospinal fluid and white matter were estimated
21 and removed using the anatomical component correction (CompCor) strategy
22 (Behzadi, Restom, Liau, & Liu, 2007). Head motion (6 motion parameters and 6

first-order temporal derivatives) as well as the main effect of task were also regressed out. The data were high-pass filtered at 0.008 Hz to preserve task-relevant high-frequency signals, which have been found to yield stronger and more reliable evidence of effects of age (Geerligs, Tsvetanov, Cam, & Henson, 2017; X. Liu et al., 2018). Pearson's correlation coefficients between each pair of regional time series were computed and transformed into Fisher's z scores. Following this procedure, undirected and weighted 264×264 FC matrices were constructed for the AVcon and AVincon conditions for each participant (Vatansever et al., 2015).

Network-based statistics

The network-based statistic (NBS) approach was used to identify functional networks underlying the differences in audiovisual integration for reading between CH and AD (Zalesky, Cocchi, Fornito, Murray, & Bullmore, 2012; Zalesky, Fornito, & Bullmore, 2010), which has been widely used in neurodevelopment research (Cignetti et al., 2018; Doesburg et al., 2016; Grayson et al., 2014). In the present study, we proceeded to detect significant nonzero connections [false discovery rate (FDR) corrected $p < 0.05$] in FC matrices for each group and condition by performing a one-sample t-test in GRETNA (<http://www.nitrc.org/projects/gretna/>) (J. Wang et al., 2015). Then, a binary matrix was created by performing a union of significant nonzero connections from the AVcon and AVincon FC matrices. The FC matrices masked by the 'union' binary matrix was inputted into the NBS to identify the significant audiovisual integration networks in the CH and AD groups, respectively. The 'union' binary

matrix was used as a mask in order to keep the same edges to be applied for the statistical comparisons (Jiang et al., 2013; Wagner et al., 2019). A less constrained primary threshold of $p < 0.05$ was used in the analysis to retain more functional connectivity information for the edges of functional networks underlying the neural differences between CH and AD. A set of supra-threshold connections were defined based on the primary threshold, which were used to determine topological components. A component (i.e. a subnetwork) is a connected graph, for which a path can be found between any two nodes. Following that, nonparametric permutation tests (5000 permutation, family-wise error rate (FWER) corrected $p < 0.05$) were performed to estimate the significance of each subnetwork based on their intensities (the sum of test statistic values across all connections). The corrected p value for a subnetwork of a given size was calculated as the proportion of permutations for which the largest component was the same size or greater. Hubs of the identified functional network for audiovisual integration were defined as those nodes whose strength was 1.5 SD (standard deviation) greater than the mean strength across all nodes in the network (X. Liu et al., 2018). Node strength is analogous to node degree in weighted networks and is defined as the sum of edge weights (i.e. Fisher's z scores) attached to a node (Fornito, Zalesky, & Bullmore, 2016; Paraskevopoulos et al., 2015).

As a final step, a 2 (group: CH and AD) \times 2 (audio-visual congruency: AVcon and AVincon) repeated measures analysis of covariance (ANCOVA) was conducted in NBS with FC matrices masked by a binary matrix creating by the union of the

significant networks involved in audiovisual integration in each group obtained in the previous step. Since the effect of head motion may confound between-group differences in functional connectivity (Geerligs et al., 2017; Siegel et al., 2017; Zeng et al., 2014), average frame wise displacement (FD) estimated based on the six head movement parameters (Power, Barnes, Snyder, Schlaggar, & Petersen, 2012) was calculated and included as a covariate. In addition, since a previous study reported gender differences in audiovisual integration processing (Ross, Del Bene, Molholm, Frey, & Foxe, 2015), gender was also taken as a covariate to control its potential effects. A primary threshold of $p < 0.01$ was applied to ANCOVA. and subnetworks with a FWER-corrected $p < 0.05$ were retained (5000 permutations). The visualization of functional networks was performed using BrainNet Viewer (Xia, Wang, & He, 2013).

Brain-behavior correlation analyses

In order to examine the relationship between reading ability and the functional networks that differed across groups, a correlation analysis was conducted between the functional networks (the congruency effect of connectivity strength) and reading performance (reading accuracy, reading fluency and phonological awareness) in CH and AD groups, respectively. In addition, to clarify the role of the functional networks during a lexical decision with congruent and incongruent stimuli, we performed a correlation analysis between the functional networks (the congruency effect of connectivity strength) and in-scanner behavior responses (accuracy and reaction time).

1 The significance level was set at $p < 0.05$ after FDR correction for multiple
2 comparisons.

3

4 **Validation analysis**

5 To evaluate the robustness of the observed functional networks underpinning the
6 differences in audiovisual integration between CH and AD, three validation
7 procedures were performed: 1) using a more stringent primary threshold of $p < 0.005$;
8 2) using another estimation method -- NBS extent (the total number of connections
9 within a component) and ; 3) using an alternative 200 ROI atlas created by Craddock
10 et al. (Craddock, James, Holtzheimer, Hu, & Mayberg, 2012).

11

12 **Functional network analysis in children with developmental dyslexia**

13 Given that audiovisual integration deficits in dyslexia may be caused by anomalies in
14 neural development, we examined whether the functional networks that differed
15 between CH and AD were disrupted in children with dyslexia. Specifically, for each
16 functional network with significant differences between CH and AD, a binary matrix
17 was first generated from the functional network. Each element of the binary matrix
18 was set to 1 if it corresponded to a nonzero element (i.e. an edge) in the functional
19 network; otherwise it was set to 0. Then, we multiplied the corresponding elements of
20 the binary matrix and the FC matrix of DD in the AVcon/AVincon condition created in
21 CONN. Thus, we obtained a new FC matrix of the DD group that was masked by the
22 binary matrix. The sum of all elements of the triangles above or below the diagonal of

the resulting matrix was taken as the network connectivity strength of the DD group. Finally, we performed a 2 (group: CH and AD) \times 2 (audio-visual congruency: AVcon and AVincon) ANCOVA of the connectivity strength of the functional network. In addition to gender, Raven IQ scores were also entered as a covariate in the analysis, since there was a significant difference in Raven IQ scores between the CH and DD groups [$t(28) = 2.16, p = 0.04$]. The significance level was set at $p < 0.05$ after FDR correction for multiple comparisons.

Results

Behavioral performances

In-scanner behavioral data for one CH and three AD were not recorded due to technical reasons. In addition, the data of another three AD were excluded because their response accuracy was too low (< 0.73), and were shown to be outliers by the boxplot (Schwertman, Owens, & Adnan, 2004) in both the AVcon and AVincon conditions. Accordingly, the behavioral performance results were based on the remaining data of 17 CH and 15 AD.

The average accuracy in the AVcon and AVincon conditions was 0.89 (SD = 0.11) and 0.76 (SD = 0.19) for CH, and was 0.98 (SD = 0.03) and 0.95 (SD = 0.05) for AD, respectively. Wilcoxon signed rank tests showed that the accuracy in the AVcon condition was higher than that in the AVincon condition for both CH ($p = 0.009$) and AD ($p = 0.029$). Mann-Whitney tests showed that the accuracy of AD was higher than

that of CH in both AVcon ($p < 0.001$) and AVincon ($p < 0.001$) conditions.

The average reaction time in the AVcon and AVincon conditions was 744.03 ms (SD = 68.12 ms) and 799.61 ms (SD = 84.59 ms) for CH, and was 616.66 ms (SD = 76.13 ms) and 662.50 ms (SD = 83.82 ms) for AD. A 2 (group: CH and AD) \times 2 (audio-visual congruency: AVcon and AVincon) ANCOVA with gender as a covariate revealed a significant main effect of group [$F(1, 29) = 22.37, p < 0.001$, partial $\eta^2 = 0.44$]. CH showed longer reaction time than AD. The main effect of audio-visual congruency was near-significant [$F(1, 29) = 3.75, p = 0.063$], but the interaction between group and audio-visual congruency [$F(1, 29) = 0.33, p = 0.570$] was not significant.

NBS analysis results

The NBS analysis revealed a large-scale functional network (173 nodes and 273 edges) for the congruency effect (AVcon > AVincon) in AD, mainly encompassing intra-regional connectivity within the prefrontal, occipital and limbic cortices, as well as inter-regional connectivity between the prefrontal and temporal cortices, between the prefrontal and parietal cortices and between the temporal and occipital cortices. The hubs included the left STG, the right MTG, the left lingual gyrus (LG), the left cuneus, the right middle occipital gyrus (MOG), the right supramarginal gyrus (SMG), the right insula, the right precuneus, the right CG and the right parahippocampal gyrus (PHG) (Figure 1). Additionally, a functional network (188 nodes and 280 edges) for

1 the incongruency effect ($AV_{incon} > AV_{con}$) was detected in AD, mainly including the
2 intra-regional connectivity within the prefrontal, occipital, parietal cortices and the
3 motor strip, and inter-regional connectivity between the prefrontal and parietal
4 cortices and between the parietal cortex and the motor strip. The hubs were the
5 bilateral STG, the bilateral IPL, the bilateral postcentral gyrus (PostCG), the left
6 precuneus, the left precentral gyrus (PreCG), the right medial frontal gyrus
7 (MedialFG), the left insula and the bilateral CG (Figure 1).

8

9 However, the contrast of $AV_{con} > AV_{incon}$ failed to identify a significant functional
10 network in CH. The incongruency effect ($AV_{incon} > AV_{con}$) was detected in a
11 functional network (222 nodes and 368 edges) in CH primarily encompassing
12 intra-regional connectivity within the prefrontal, occipital, parietal cortices and motor
13 strip, and inter-regional connectivity between the prefrontal and limbic cortices. The
14 hubs were the bilateral LG, the bilateral cuneus, the right FuG, the left MOG, the right
15 superior/middle/inferior frontal gyrus (SFG/MFG/IFG), the bilateral inferior parietal
16 gyrus (IPL), the left PreCG, the bilateral lentiform nucleus (LN) and the right insula
17 (Figure 1).

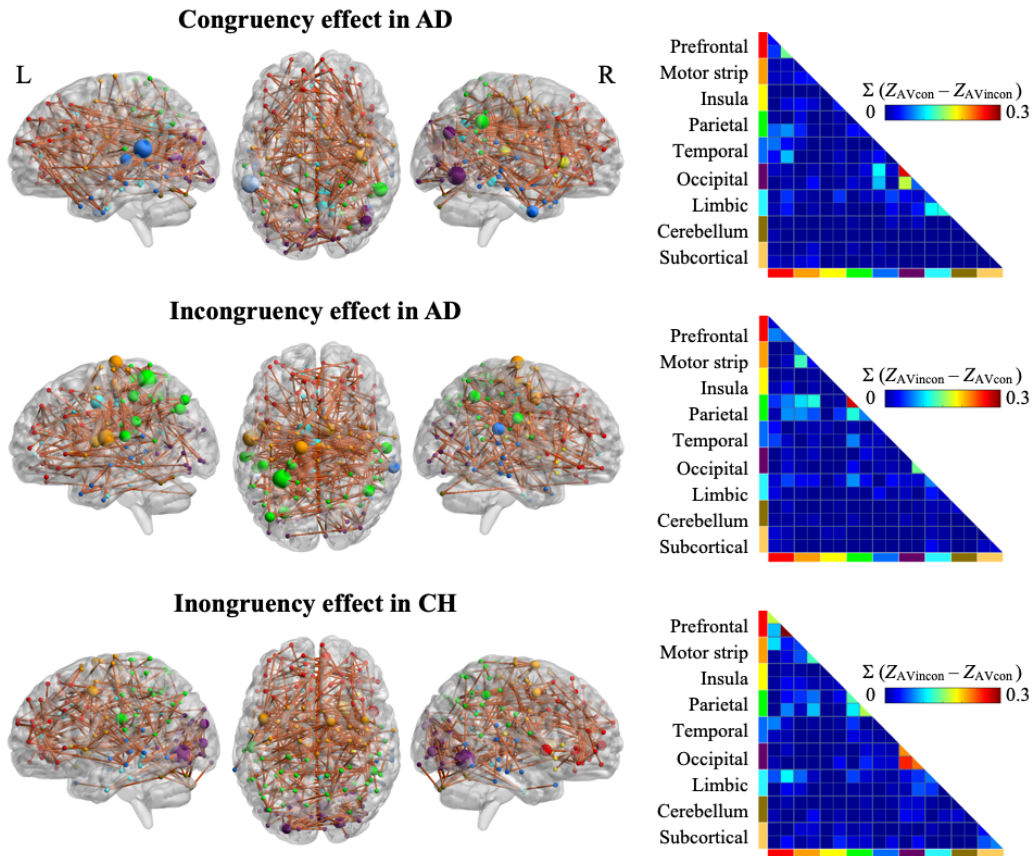


Figure 1 The functional networks underpinning congruency effect ($AV_{con} > AV_{incon}$) or incongruency effect ($AV_{incon} > AV_{con}$) of audiovisual integration in normally developing children and adults, respectively. The colors of the nodes in the brain plots indicate the lobe (coded by color bands along the matrix plots) to which they belong. The large nodes represent hubs, whose sizes are proportional to the node strengths. The matrix plots in the right panel represent connectivity strength between pairs of the 9 brain lobes. Within each lobe, left hemisphere nodes are at the top (left) while right hemisphere nodes are at the bottom (right), separated by thin lines. The color of each element in the matrices represent the sum of the weight of all the edges for the connected lobes. CH = normally developing children, AD = adults. AV_{con} = audiovisually congruent characters, AV_{incon} = audiovisually incongruent characters.

1 L = left, R = right.

2

3 Critically, the 2 (group: CH and AD) \times 2 (audio-visual congruency: AVcon and
4 AVincon) repeated measures ANCOVA identified a significant interaction between
5 group and audio-visual congruency in a functional network comprising 5 nodes and 3
6 edges, which could be segmented into two subnetworks. The first subnetwork
7 encompassed the left STG, the right MedialFG and the right SFG, forming a
8 prefrontal-superior temporal functional network [interaction effect: $F(1, 32) = 21.65$,
9 $p < 0.001$, partial $\eta^2 = 0.40$]. The second subnetwork comprised the left thalamus and
10 the right lentiform nucleus [interaction effect: $F(1, 32) = 15.75$, $p < 0.001$, partial $\eta^2 =$
11 0.33]. In both subnetworks, the AD group showed a significant congruency effect
12 (AVcon > AVincon) [first subnetwork: $F(1, 32) = 22.17$, $p < 0.001$, partial $\eta^2 = 0.41$;
13 second subnetwork: $F(1, 32) = 5.73$, $p = 0.023$, partial $\eta^2 = 0.15$], while CH showed
14 an incongruency effect (AVcon < AVincon) [first subnetwork: $F(1, 32) = 5.10$, $p =$
15 0.031 , partial $\eta^2 = 0.14$; second subnetwork: $F(1, 32) = 12.60$, $p = 0.001$, partial $\eta^2 =$
16 0.28] (Figure 2A).

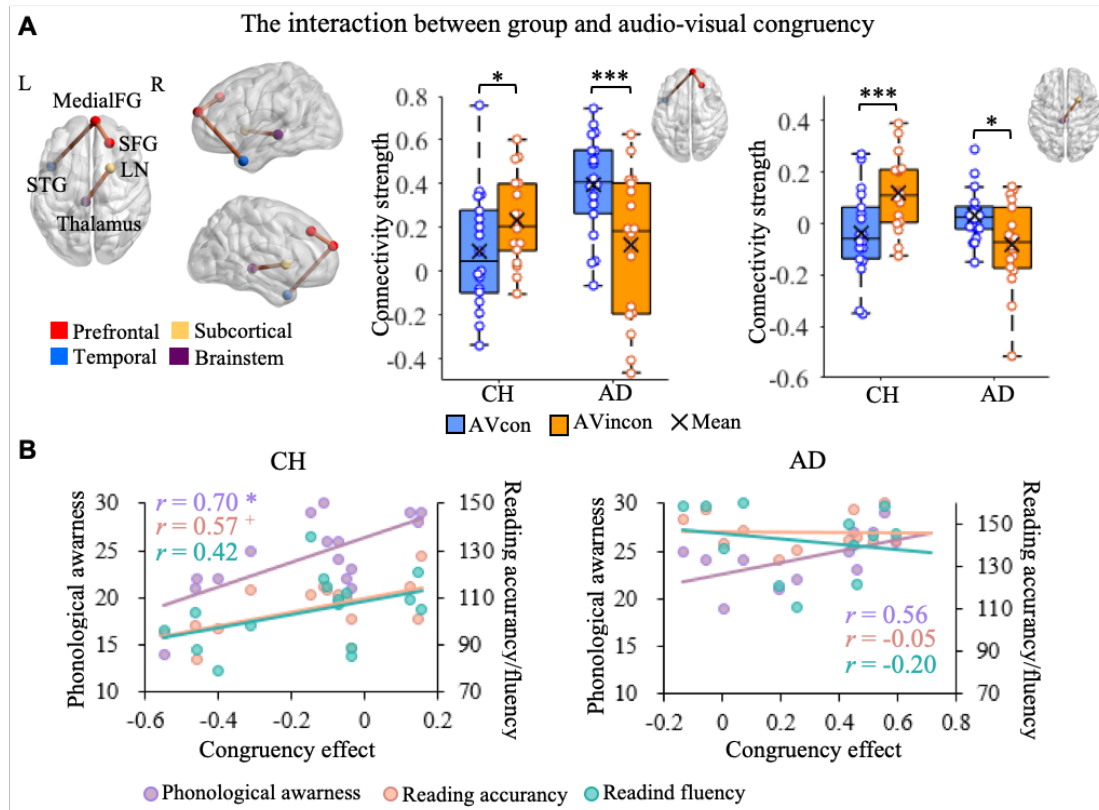


Figure 2 Functional networks underlying differences in audiovisual integration for reading between children and adults, and their correlations with reading ability. (A) Functional networks with significant interactions between group and audio-visual congruency. The color of the nodes indicates the lobes they belong to. (B) Correlation between the congruency effect (calculated by connectivity strength of the prefrontal-superior temporal network) and reading performance in normally developing children and adults. CH = normally developing children, AD = adults. AVcon = audiovisually congruent characters, AVincon = audiovisually incongruent characters. MedialFG = medial frontal gyrus, SFG = superior frontal gyrus, STG = superior temporal gyrus, LN = lentiform nucleus. L = left, R = right. + $P < 0.1$, * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$.

Correlation analyses revealed that in CH, the congruency effect (calculated by connectivity strength of the prefrontal-superior temporal network) correlated with reading accuracy ($r = 0.57$, FDR-corrected $p = 0.06$) and phonological awareness ($r = 0.70$, FDR-corrected $p = 0.018$), but not with reading fluency ($r = 0.42$, FDR-corrected $p = 0.162$). There were no significant correlations in the AD group (reading accuracy: $r = -0.05$, FDR-corrected $p = 0.883$; reading fluency: $r = -0.20$, FDR-corrected $p = 0.646$; phonological awareness: $r = 0.56$, FDR-corrected $p = 0.116$) (Figure 2B). A Spearman correlation analysis showed that in the AD group, the congruency effect (calculated by connectivity strength of the prefrontal-superior temporal network) was positively correlated with accuracy in the AVincon condition ($r = 0.55$, uncorrected $p = 0.032$), while in the group of CH, the correlation showed a similar trend, but was not significant ($r = 0.39$, uncorrected $p = 0.125$).

Validation results

The prefrontal-superior temporal network was repeated by the previous NBS procedure with a more stringent primary threshold ($p < 0.005$), and an additional estimation method based on the NBS extent with a primary threshold of $p < 0.01$. However, the analysis using a $p < 0.005$ threshold revealed no connectivity between the right SFG and the right MedialFG, and the analysis based on the NBS extent revealed no connectivity between the left thalamus and the right lentiform nucleus (Figure S1A). Using an alternative Craddock atlas, NBS analysis revealed a significant interaction between group and audio-visual congruency in a functional

network involving the left MFG, the right frontal pole, the bilateral planum temporale/STG and the left lateral occipital cortex [interaction effect: $F(1, 32) = 11.76$, $p = 0.002$, partial $\eta^2 = 0.27$; AD showed a significant congruency effect: $F(1, 32) = 11.82$, $p = 0.002$, partial $\eta^2 = 0.27$; CH showed a near-significant incongruency effect: $F(1, 32) = 2.88$, $p = 0.099$, partial $\eta^2 = 0.08$] (Figure S1B). The validation results highlighted the reproducibility of the identified prefrontal-superior temporal network underpinning the differences in audiovisual integration between children and adults.

Results of functional network analysis in children with dyslexia

Since only the prefrontal-superior temporal network was identified in validation analysis, we examined this brain network in dyslexic group. ANCOVA revealed a marginal significant interaction between group and audiovisual congruency [$F(1, 26) = 3.06$, $p = 0.092$, partial $\eta^2 = 0.11$], but no significant main effects of group [$F(1, 26) = 1.10$, $p = 0.305$] or audiovisual congruency [$F(1, 26) = 1.53$, $p = 0.227$] were found.

A simple effects analysis showed that there was a significant incongruency effect (AVincon > AVcon) in CH [$F(1, 26) = 5.58$, $p = 0.026$, partial $\eta^2 = 0.18$], but not in DD [$F(1, 26) = 0.20$, $p = 0.661$] (Figure 3).

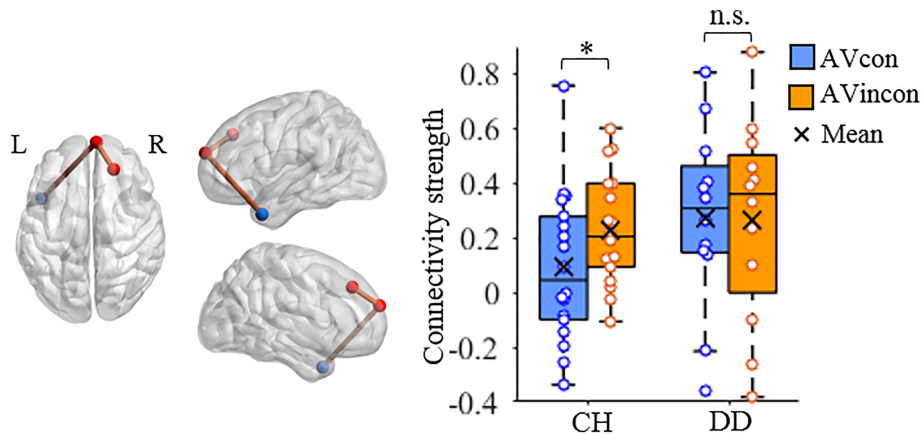


Figure 3 Connectivity strength of normally developing children and children with dyslexia in the prefrontal-superior temporal network. CH = normally developing children, DD = children with dyslexia. AVcon = audiovisually congruent characters, AVincon = audiovisually incongruent characters. L = left, R = right. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$. n.s. = not significant.

Discussion

The present study aimed to explore the differences in functional brain networks in audiovisual integration for reading between children and adults. We found that during the lexical decision task, adults showed greater connectivity than children in a prefrontal-superior temporal network (encompassing the right medial frontal gyrus, the right superior frontal gyrus and the left superior temporal gyrus) and a thalamus-lentiform nucleus network (encompassing the left thalamus and the right lentiform nucleus), suggesting that these networks are associated with the development of audiovisual integration for reading. Moreover, the prefrontal-superior temporal network was found to be disrupted in children with dyslexia, thus

confirming its role in audiovisual integration for reading. Taken together, our findings reveal, for the first time, the brain mechanisms of audiovisual integration for reading in adults and children, as part of multimodal information processing in higher cognition.

Functional networks of audiovisual integration for reading

Much research has identified areas associated with audiovisual integration such as STG, MTG, MFG, IFG, FuG, LG, cuneus, IPL, precuneus, insula and CG (Blau et al., 2008; Erickson, Heeg, Rauschecker, & Turkeltaub, 2014; Hocking & Price, 2009; Holloway et al., 2015; Raij et al., 2000; van Atteveldt et al., 2010; van Atteveldt et al., 2006; van Atteveldt et al., 2004; W. Xu et al., 2019). To our knowledge, the present study is the first to identify the functional brain networks of audiovisual integration for reading in both child and adult readers. Consistent with our predictions, audiovisual integration recruited a large-scale functional network in adults, involving intra-regional connectivity within the occipital cortex and inter-regional connectivity between the temporal and occipital cortices and between the prefrontal and temporal cortices. The left posterior STG was identified as the main hub, in line with previous findings showing its core role in audiovisual integration for both speech and non-speech stimuli (Erickson et al., 2014; Bethany Plakke & Romanski, 2019; van Atteveldt et al., 2004; Ye, Rüsseler, Gerth, & Münte, 2017). When presented with incongruent audiovisual stimuli during a lexical decision task, both children and adults showed enhanced intra-regional connectivity within the prefrontal, occipital,

parietal regions, as well as the motor strip, consistent with evidence implicating frontoparietal and pre-supplementary motor areas in response inhibition (R. Zhang, Geng, & Lee, 2017). A previous neural framework for reading (Price & Devlin, 2011) has proposed that connectivity within the occipital cortex reflects bottom-up transmission of visual features to ventral occipitotemporal cortex (vOT, including FuG and LG), and the connectivity between the STG and vOT reflects top-down generation of predictions formed from prior experience. Moreover, top-down processing is modulated by higher order regions (prefrontal cortex) associated with attention and task demands (E. K. Miller & J. D. Cohen, 2001; Price & Devlin, 2011; Yoncheva, Zevin, Maurer, & McCandliss, 2010). In the current lexical decision task, the input of auditory speech sounds might strengthen the interaction between top-down and bottom-up hierarchies, which in turn affect the recognition of visual characters. However, we did not detect significant networks in the AVcon > AVincon contrast in normally developing children. Presumably, children might be less sensitive to the audiovisual congruency effect during reading, given a lower level of linguistic knowledge and experience.

Functional networks underlying the developmental changes in audiovisual integration for reading

Network analysis revealed that in a prefrontal-superior temporal network, adults showed stronger connectivity in the AVcon condition than in the AVincon condition (congruency effect), while children showed a reverse pattern. The congruency effect

(calculated by the connectivity strength of the network) was positively correlated with reading accuracy and phonological awareness in children, but not in adults. This result suggests that this brain network is vital to reading skill, which is more pronounced in developing readers. Indeed, previous studies have highlighted the role of phonology in reading (Ho & Bryant, 1997; Karipidis et al., 2017; Melby-Lervag, Lyster, & Hulme, 2012), and additionally, reliance on phonological processing is greater in children than in adults (X. Liu et al., 2018). Two of the nodes identified in the network are the MedialFG and SFG and are located in the PFC (Carlen, 2017), which is part of the associative cortex of the frontal lobe (Calvert, 2001; Fuster, 1985). The PFC receives a wide array of sensory inputs from multiple modalities (Macaluso & Driver, 2005; MartÃ-nez-Sanchis, 2014; Bethany Plakke & Romanski, 2019; Sugihara, Diltz, Averbek, & Romanski, 2006). In addition, the PFC has been hypothesized to support executive function, such as attention, inhibitory control and decision making (Carlen, 2017). Specifically, the bilateral MedialFG is generally involved in conflict detection (Aarts, Roelofs, & van Turenout, 2009; Bolger, Hornickel, Cone, Burman, & Booth, 2008; Doehrmann & Naumer, 2008; Noppeney, Ostwald, & Werner, 2010) and attentional control (Aarts et al., 2009; Bush, Luu, & Posner, 2000; Earl K. Miller & Jonathan D. Cohen, 2001). Moreover, the right SFG has been reported to subserve conflict resolution mechanisms by utilizing top-down control of attentional resources (Corbetta & Shulman, 2002; Muller et al., 2011; R. Zhang et al., 2017). Besides Medial FG and SFG, another critical node that was shown to be part of the network is the left anterior STG. This finding is consistent

1 with previous studies reporting the engagement of the anterior superior temporal
2 cortex in letter-speech sound integration (Hocking & Price, 2009; van Atteveldt et al.,
3 2006). Furthermore, the anterior STG has been identified as a semantic hub enabling
4 activation of semantic representations, irrespective of the input modality, such as
5 written words, auditory sounds and pictures (Lambon Ralph, Sage, Jones, &
6 Mayberry, 2010; Visser & Lambon Ralph, 2011). During reading, semantic,
7 phonological and orthographic representations are automatically accessed, even when
8 semantic processing is not required (Perfetti & Tan, 1998; Y. Xu, Pollatsek, & Potter,
9 1999; S. L. Zhang, Perfetti, & Yang, 1999), and this is especially true in Chinese,
10 which has direct mappings between orthography and semantics (Y. Liu & Perfetti,
11 2003). Thus in our lexical decision task, both phonological and semantic
12 representations are activated (Specht et al., 2003; Yates, Locker, & Simpson, 2003),
13 but given that semantic skills undergo prolonged development throughout childhood
14 and adolescence (Moore-Parks et al., 2010; Vannest, Karunanayaka, Schmithorst,
15 Szaflarski, & Holland, 2009), adults might automatically activate semantic
16 representations, whereas children might tend to activate phonological representations,
17 which is reflected in differences in connectivity in the anterior STG between children
18 and adults.

19
20 There is anatomical evidence of structural projections from the STG to the PFC in
21 nonhuman primates (Petrides & Pandya, 2002; B. Plakke & Romanski, 2016) as well
22 as in humans (Garell et al., 2012). The uncinate fasciculus and the arcuate fasciculus

1 are two main fiber tracts that link the STG with the PFC, forming the anatomical
2 substrates for the transmission of auditory or multisensory information (Bethany
3 Plakke & Romanski, 2019). Functionally, the connectivity between the STG and PFC
4 has been linked to attention shifting (Pammer, Hansen, Holliday, & Cornelissen, 2006;
5 Paraskevopoulos et al., 2015) and behavioral responses to visual stimuli (Donner et al.,
6 2007). Attention is a vital factor for the late development of audiovisual integration
7 skill (Burr & Gori, 2012; Dick et al., 2010), and it has been found to mediate
8 audiovisual integration processing at multiple stages (including visual and auditory
9 processing, spatiotemporal realignment, congruency matching and semantic analysis)
10 in both bottom-up and top-down fashion (Koelewijn, Bronkhorst, & Theeuwes, 2010;
11 Navarra, Alsius, Soto-Faraco, & Spence, 2010; Talsma et al., 2010). According to the
12 framework of the multifaceted interplay between multisensory integration and
13 attention (Talsma et al., 2010), audiovisual integration tends to occur pre-attentively
14 (i.e., bottom-up) when speech inputs are congruent with visual characters, which in
15 turn enhances the perceptual processing of task-relevant modality. In contrast, when
16 speech input is in conflict with visual character input, top-down attentional
17 mechanisms are required to inhibit the task-irrelevant stimuli that act as attention
18 capturing distractors. Evidence of such top-down mechanisms comes from the
19 positive correlation between the prefrontal-superior temporal network and the
20 accuracy in the AVincon condition. Consequently, compared with developing children,
21 the maturation of the prefrontal-superior temporal network may allow skilled adult
22 readers to take more advantage of congruent phonology and suppress more

1 incongruent interference for the recognition of visually presented characters.

2

3 In addition, the differences in audiovisual integration between children and adults
4 were also found in a thalamus-lentiform nucleus network (the connectivity between
5 the left thalamus and the right lentiform nucleus). The thalamus is an interface
6 through which nearly all sensory information must pass it before reaching the cerebral
7 cortex (McCormick & Bal, 1994). The lentiform nucleus is a part of the striatum that
8 receives massive projections from thalamus (Russchen & Jonker, 1988). Previous
9 studies have demonstrated that the fronto-striato-thalamic pathway was associated
10 with inhibitory capacity, which develops from childhood to adulthood (Rubia, Smith,
11 Taylor, & Brammer, 2007). Therefore, the subcortical thalamus-lentiform nucleus
12 network may support the inhibition of interference caused by incongruent auditory
13 speech sounds. However, the result was not replicated in the validation procedure, and
14 therefore need further verification.

15

16 **The disruption of the prefrontal-superior temporal network in developmental** 17 **dyslexia**

18 To further confirm whether the prefrontal-superior temporal networks is a key neural
19 circuit for audiovisual integration specific to reading, we examined this networks in a
20 group of dyslexic children with impaired reading skills. In line with our expectation,
21 we observed disruption of the prefrontal-superior temporal network in dyslexic
22 children. This result is accordance with a previous evidence of abnormalities in a

functional network composed of STG/STS and medial prefrontal cortex in German dyslexics in an audiovisual speech integration task (Ye et al., 2017). Since the prefrontal-superior temporal network is involved in cross-modal attention shifts (Pammer et al., 2006; Paraskevopoulos et al., 2015), its disruption potentially signals difficulties in shifting attention between modalities in children with dyslexia, a phenomenon known as “sluggish attentional shifting” (Hari & Renvall, 2002; Harrar et al., 2014). As a result, Children with dyslexia may focus their limited attention resources on task-related visual stimuli (characters) and be less affected (facilitated/inhibited) by auditory input. In all, the lack of a congruency/incongruency effect in children with dyslexia revealed an atypical development of the prefrontal-superior temporal network. The present findings further elucidate the neural foundations of developmental dyslexia and thus have the potential to inform assessment and intervention programs in developmental dyslexia (Hillock-Dunn & Wallace, 2012; Schlaggar et al., 2002).

Limitations.

Some limitations should be considered in light of the present findings. First, we employed a visual lexical decision task to examine the effect of simultaneously presenting speech sounds and visual characters in a realistic reading context (Knoop-van Campen, Segers, & Verhoeven, 2020). However, further research is required in order to test whether our findings apply in the context of auditory perception, that is, the extent to which presenting visual information affects auditory

1 perception during the audiovisual integration processing.

2

3 Second, our findings apply to the Chinese writing system, which is non-transparent
4 and drastically different from alphabetic languages regarding in both the visual
5 features of the written scripts and in its orthography-to-phonology correspondences.

6 As previous studies have shown that audiovisual integration depends on the extent of
7 orthographic transparency (Holloway et al., 2015), further work is necessary to
8 examine the generalizability of the present results to transparent/semitransparent
9 writing systems.

10

11 Third, although the effective sample size of the present study meets the requirement
12 for acceptable statistical power, future studies with a larger sample of participants are
13 needed to obtain more reliable and repeatable results. In addition, further studies
14 might consider recruiting participants from a wider age range and use a longitudinal
15 design to assess changes in the networks of audiovisual integration for reading along
16 the neurodevelopment trajectory.

17

18 **Conclusion**

19 The present study revealed the differences in a prefrontal-superior temporal network
20 that is involved in audiovisual integration for reading between normal developing
21 children and skilled adult readers. The findings presumably reflect the effect of
22 attention modulation in audiovisual integration. In addition, the network identified

1 was disrupted in a group of children with developmental dyslexia, thus highlighting
2 its importance in reading. We argue that the present study is the first to unveil the
3 neural mechanisms of audiovisual integration for reading in children and adults,
4 potentially reflecting neurodevelopmental changes due to the development of reading
5 skills, and advancing our understanding of neural correlates of multimodal sensory
6 integration in humans.

7

8 **Acknowledgements**

9 We thank all the children and adults who participated in this study. This work was
10 supported by the National Natural Science Foundation of China (No. 31671155 and
11 31800954), CAS Key Laboratory of Behavioral Science, Institute of Psychology.

12

13 **Competing Interests**

14 The authors declare that they have no competing interests.

References

- Aarts, E., Roelofs, A., & van Turenout, M. (2009). Attentional control of task and response in lateral and medial frontal cortex: brain activity and reaction time distributions. *Neuropsychologia*, 47(10), 2089-2099. doi:10.1016/j.neuropsychologia.2009.03.019
- Barutchu, A., Danaher, J., Crewther, S. G., Innes-Brown, H., Shivdasani, M. N., & Paolini, A. G. (2010). Audiovisual integration in noise by children and adults. *Journal of Experimental Child Psychology*, 105(1-2), 38-50. doi:10.1016/j.jecp.2009.08.005
- Behzadi, Y., Restom, K., Liao, J., & Liu, T. T. (2007). A component based noise correction method (CompCor) for BOLD and perfusion based fMRI. *NeuroImage*, 37(1), 90-101. doi:10.1016/j.neuroimage.2007.04.042
- Beierholm, U. R., & Adams, W. J. (2016). The Development of Audio-Visual Integration for Temporal Judgements. *PLoS Computational Biology*, 12(4), e1004865. doi:10.1371/journal.pcbi.1004865
- Blau, V., Reithler, J., van Atteveldt, N., Seitz, J., Gerretsen, P., Goebel, R., & Blomert, L. (2010). Deviant processing of letters and speech sounds as proximate cause of reading failure: a functional magnetic resonance imaging study of dyslexic children. *Brain*, 133(Pt 3), 868-879. doi:10.1093/brain/awp308
- Blau, V., van Atteveldt, N., Ekkebus, M., Goebel, R., & Blomert, L. (2009). Reduced Neural Integration of Letters and Speech Sounds Links Phonological and Reading Deficits in Adult Dyslexia. *Current Biology*, 19(6), 503-508. doi:10.1016/j.cub.2009.01.065
- Blau, V., van Atteveldt, N., Formisano, E., Goebel, R., & Blomert, L. (2008). Task-irrelevant visual letters interact with the processing of speech sounds in heteromodal and unimodal cortex. *European Journal of Neuroscience*, 28(3), 500-509. doi:10.1111/j.1460-9568.2008.06350.x
- Blomert, L., & Vaessen, A. (2009). Differentiaal Diagnostiek van Dyslexie: Cognitieve analyse van lezen en spellen [Dyslexia Differential Diagnosis: Cognitive analysis of reading and spelling]. Amsterdam: Boom Test Publishers.
- Bolger, D. J., Hornickel, J., Cone, N. E., Burman, D. D., & Booth, J. R. (2008). Neural correlates of orthographic and phonological consistency effects in children. *Hum Brain Mapp*, 29(12), 1416-1429. doi:10.1002/hbm.20476
- Burr, D., & Gori, M. (2012). *Multisensory Integration Develops Late in Humans*: PubMed.
- Bush, G., Luu, P., & Posner, M. I. (2000). Cognitive and emotional influences in anterior cingulate cortex. *Trends in Cognitive Sciences*, 4(6), 215-222. doi:10.1016/s1364-6613(00)01483-2
- Calvert, G. A. (2001). Crossmodal processing in the human brain: Insights from functional neuroimaging studies. *Cerebral Cortex*, 11(12), 1110-1123. doi:10.1093/cercor/11.12.1110
- Carlen, M. (2017). What constitutes the prefrontal cortex? *Science*, 358(6362), 478-482. doi:10.1126/science.aan8868
- Cignetti, F., Vaugoyeau, M., Decker, L. M., Grosbras, M.-H., Girard, N., Chaix, Y., Péran, P., & Assaiante, C. (2018). Brain network connectivity associated with anticipatory postural control in children and adults. *Cortex*, 108, 210-221. doi:10.1016/j.cortex.2018.08.013
- Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nat Rev Neurosci*, 3(3), 201-215. doi:10.1038/nrn755
- Craddock, R. C., James, G. A., Holtzheimer, P. E., 3rd, Hu, X. P., & Mayberg, H. S. (2012). A whole brain fMRI atlas generated via spatially constrained spectral clustering. *Hum Brain Mapp*,

- 33(8), 1914-1928. doi:10.1002/hbm.21333
- Dick, A. S., Solodkin, A., & Small, S. L. (2010). Neural development of networks for audiovisual speech comprehension. *Brain Lang*, 114(2), 101-114. doi:10.1016/j.bandl.2009.08.005
- Doehrmann, O., & Naumer, M. J. (2008). Semantics and the multisensory brain: How meaning modulates processes of audio-visual integration. *Brain Research*, 1242, 136-150. doi:10.1016/j.brainres.2008.03.071
- Doesburg, S. M., Tingling, K., MacDonald, M. J., & Pang, E. W. (2016). Development of Network Synchronization Predicts Language Abilities. *Journal of Cognitive Neuroscience*, 28(1), 55-68. doi:10.1162/jocn_a_00879
- Donner, T. H., Siegel, M., Oostenveld, R., Fries, P., Bauer, M., & Engel, A. K. (2007). Population activity in the human dorsal pathway predicts the accuracy of visual motion detection. *Journal of Neurophysiology*, 98(1), 345-359. doi:10.1152/jn.01141.2006
- Driver, J., & Noesselt, T. (2008). Multisensory Interplay Reveals Crossmodal Influences on 'Sensory-Specific' Brain Regions, Neural Responses, and Judgments. *Neuron*, 57(1), 11-23. doi:10.1016/j.neuron.2007.12.013
- Erickson, L. C., Heeg, E., Rauschecker, J. P., & Turkeltaub, P. E. (2014). An ALE meta-analysis on the audiovisual integration of speech signals. *Human Brain Mapping*, 35(11), 5587-5605. doi:10.1002/hbm.22572
- Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39(2), 175-191. doi:10.3758/bf03193146
- Fornito, A., Zalesky, A., & Bullmore, E. T. (2016). Fundamentals of Brain Network Analysis. 115-136.
- Froyen, D., Bonte, M., van Atteveldt, N., & Blomert, L. (2009). The Long Road to Automation: Neurocognitive Development of Letter-Speech Sound Processing. *Journal of Cognitive Neuroscience*, 21(3), 567-580. doi:10.1162/jocn.2009.21061
- Fuster, J. M. (1985). *The Prefrontal Cortex and Temporal Integration*: Springer US.
- Garell, P. C., Bakken, H., Greenlee, J. D. W., Volkov, I., Reale, R. A., Oya, H., Kawasaki, H., Howard, M. A., & Brugge, J. F. (2012). Functional Connection Between Posterior Superior Temporal Gyrus and Ventrolateral Prefrontal Cortex in Human. *Cerebral Cortex*, 23(10), 2309-2321. doi:10.1093/cercor/bhs220
- Geerligs, L., Tsvetanov, K. A., Cam, C. A. N., & Henson, R. N. (2017). Challenges in measuring individual differences in functional connectivity using fMRI: The case of healthy aging. *Human Brain Mapping*, 38(8), 4125-4156. doi:10.1002/hbm.23653
- Grayson, D. S., Ray, S., Carpenter, S., Iyer, S., Dias, T. G., Stevens, C., Nigg, J. T., & Fair, D. A. (2014). Structural and functional rich club organization of the brain in children and adults. *PLoS ONE*, 9(2), e88297. doi:10.1371/journal.pone.0088297
- Hari, R., & Renvall, H. (2002). Impaired processing of rapid stimulus sequences in dyslexia. *Trends in Cognitive Sciences*, 5(12), 525-532.
- Harrar, V., Tammam, J., Perez-Bellido, A., Pitt, A., Stein, J., & Spence, C. (2014). Multisensory integration and attention in developmental dyslexia. *Curr Biol*, 24(5), 531-535. doi:10.1016/j.cub.2014.01.029
- Hillock-Dunn, A., & Wallace, M. T. (2012). Developmental changes in the multisensory temporal binding window persist into adolescence. *Developmental Science*, 15(5), 688-696. doi:10.1111/j.1467-7687.2012.01171.x

- Ho, C. S.-H., & Bryant, P. (1997). Phonological skills are important in learning to read Chinese. *Developmental Psychology*, 33(6), 946-951. doi:10.1037/0012-1649.33.6.946
- Hocking, J., & Price, C. J. (2009). Dissociating verbal and nonverbal audiovisual object processing. *Brain & Language*, 108(2), 89-96. doi:10.1016/j.bandl.2008.10.005
- Holloway, I. D., van Atteveldt, N., Blomert, L., & Ansari, D. (2015). Orthographic dependency in the neural correlates of reading: evidence from audiovisual integration in English readers. *Cereb Cortex*, 25(6), 1544-1553. doi:10.1093/cercor/bht347
- Jiang, G., Wen, X., Qiu, Y., Zhang, R., Wang, J., Li, M., Ma, X., Tian, J., & Huang, R. (2013). Disrupted topological organization in whole-brain functional networks of heroin-dependent individuals: a resting-state fMRI study. *PLoS One*, 8(12), e82715.
- Karipidis, I. I., Pleisch, G., Rothlisberger, M., Hofstetter, C., Dornbierer, D., Stampfli, P., & Brem, S. (2017). Neural initialization of audiovisual integration in prereaders at varying risk for developmental dyslexia. *Hum Brain Mapp*, 38(2), 1038-1055. doi:10.1002/hbm.23437
- Kast, M., Bezzola, L., Jäncke, L., & Meyer, M. (2011). Multi- and unisensory decoding of words and nonwords result in differential brain responses in dyslexic and nondyslexic adults. *Brain Lang*, 119(3), 136-148. doi:10.1016/j.bandl.2011.04.002
- Knoop-van Campen, C. A. N., Segers, E., & Verhoeven, L. (2020). Effects of audio support on multimedia learning processes and outcomes in students with dyslexia. *Computers & Education*, 150. doi:10.1016/j.compedu.2020.103858
- Koelewijn, T., Bronkhorst, A., & Theeuwes, J. (2010). Attention and the multiple stages of multisensory integration: A review of audiovisual studies. *Acta Psychol (Amst)*, 134(3), 372-384. doi:10.1016/j.actpsy.2010.03.010
- Lambon Ralph, M. A., Sage, K., Jones, R. W., & Mayberry, E. J. (2010). Coherent concepts are computed in the anterior temporal lobes. *Proc Natl Acad Sci U S A*, 107(6), 2717-2722. doi:10.1073/pnas.0907307107
- Li, D., Chen, G., & Jin, Y. (1989). Combined Raven's Matrices Test (CRT): Handbook of Chinese (revised.). Shanghai, China: East China Normal University Press (in Chinese).
- Liu, X., Gao, Y., Di, Q., Hu, J., Lu, C., Nan, Y., Booth, J. R., & Liu, L. (2018). Differences between child and adult large-scale functional brain networks for reading tasks. *Human Brain Mapping*, 39(2), 662-679. doi:10.1002/hbm.23871
- Liu, Y., & Perfetti, C. A. (2003). The time course of brain activity in reading English and Chinese: an ERP study of Chinese bilinguals. *Human Brain Mapping*, 18(3), 167-175. doi:10.1002/hbm.10090
- Macaluso, E., & Driver, J. (2005). Multisensory spatial interactions: a window onto functional integration in the human brain. *Trends Neurosci*, 28(5), 264-271. doi:10.1016/j.tins.2005.03.008
- Martínez-Sánchez, S. (2014). Neurobiological foundations of multisensory integration in people with autism spectrum disorders: the role of the medial prefrontal cortex. *Frontiers in Human Neuroscience*, 8, 970. doi:10.3389/fnhum.2014.00970
- Mccormick, D. A., & Bal, T. (1994). Sensory gating mechanisms of the thalamus. *Current Opinion in Neurobiology*, 4(4), 550-556.
- McNorgan, C., Randazzo-Wagner, M., & Booth, J. R. (2013). Cross-modal integration in the brain is related to phonological awareness only in typical readers, not in those with reading difficulty. *Front Hum Neurosci*, 7, 388. doi:10.3389/fnhum.2013.00388

- Melby-Lervag, M., Lyster, S. A., & Hulme, C. (2012). Phonological skills and their role in learning to read: a meta-analytic review. *Psychological Bulletin*, 138(2), 322-352. doi:10.1037/a0026744
- Meunier, D., Achard, S., Morcom, A., & Bullmore, E. (2009). Age-related changes in modular organization of human brain functional networks. *Neuroimage*, 44(3), 715-723. doi:10.1016/j.neuroimage.2008.09.062
- Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience*, 24(1), 167-202. doi:10.1146/annurev.neuro.24.1.167
- Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annu Rev Neurosci*, 24, 167-202. doi:10.1146/annurev.neuro.24.1.167
- Moore-Parks, E. N., Burns, E. L., Bazzill, R., Levy, S., Posada, V., & Muller, R. A. (2010). An fMRI study of sentence-embedded lexical-semantic decision in children and adults. *Brain Lang*, 114(2), 90-100. doi:10.1016/j.bandl.2010.03.009
- Muller, V. I., Habel, U., Derntl, B., Schneider, F., Zilles, K., Turetsky, B. I., & Eickhoff, S. B. (2011). Incongruence effects in crossmodal emotional integration. *NeuroImage*, 54(3), 2257-2266. doi:10.1016/j.neuroimage.2010.10.047
- Mumford, J. A. (2012). A power calculation guide for fMRI studies. *Social Cognitive and Affective Neuroscience*, 7(6), 738-742. doi:10.1093/scan/nss059
- Murray, M. M., & Wallace, M. T. (2012). The Neural Bases of Multisensory Processes. (Boca Raton, FL: CRC Press). doi:10.1201/9781439812174
- Navarra, J., Alsius, A., Soto-Faraco, S., & Spence, C. (2010). Assessing the role of attention in the audiovisual integration of speech. *Information Fusion*, 11(1), 4-11. doi:10.1016/j.inffus.2009.04.001
- Noppeney, U., Ostwald, D., & Werner, S. (2010). Perceptual Decisions Formed by Accumulation of Audiovisual Evidence in Prefrontal Cortex. *Journal of Neuroscience*, 30(21), 7434-7446. doi:10.1523/jneurosci.0455-10.2010
- Pammer, K., Hansen, P., Holliday, I., & Cornelissen, P. (2006). Attentional shifting and the role of the dorsal pathway in visual word recognition. *Neuropsychologia*, 44(14), 2926-2936. doi:10.1016/j.neuropsychologia.2006.06.028
- Paraskevopoulos, E., Kraneburg, A., Herholz, S. C., Bamidis, P. D., & Pantev, C. (2015). Musical expertise is related to altered functional connectivity during audiovisual integration. *Proc Natl Acad Sci U S A*, 112(40), 12522-12527. doi:10.1073/pnas.1510662112
- Perfetti, C. A., & Tan, L. H. (1998). The time course of graphic, phonological, and semantic activation in Chinese character identification. *Journal of Experimental Psychology-Learning Memory and Cognition*, 24(1), 101-118. doi:10.1037/0278-7393.24.1.101
- Petrides, M., & Pandya, D. N. (2002). Comparative cytoarchitectonic analysis of the human and the macaque ventrolateral prefrontal cortex and corticocortical connection patterns in the monkey. *European Journal of Neuroscience*, 16, 291-310.
- Plakke, B., & Romanski, L. M. (2016). Neural circuits in auditory and audiovisual memory. *Brain Res*, 1640(Pt B), 278-288. doi:10.1016/j.brainres.2015.11.042
- Plakke, B., & Romanski, L. M. (2019). *Audiovisual Integration in the Primate Prefrontal Cortex: The Auditory Perspective: Multisensory Processes*.
- Power, J. D., Barnes, K. A., Snyder, A. Z., Schlaggar, B. L., & Petersen, S. E. (2012). Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. *NeuroImage*, 59(3), 2142-2154. doi:10.1016/j.neuroimage.2011.10.018

- Power, J. D., Cohen, A. L., Nelson, S. M., Wig, G. S., Barnes, K. A., Church, J. A., Vogel, A. C., Laumann, T. O., Miezin, F. M., Schlaggar, B. L., & Petersen, S. E. (2011). Functional network organization of the human brain. *Neuron*, 72(4), 665-678. doi:10.1016/j.neuron.2011.09.006
- Price, C. J., & Devlin, J. T. (2011). The interactive account of ventral occipitotemporal contributions to reading. *Trends Cogn Sci*, 15(6), 246-253. doi:10.1016/j.tics.2011.04.001
- Raij, T., Uutela, K., & Hari, R. (2000). Audiovisual integration of letters in the human brain. *Neuron*, 28(2), 617-625. doi:10.1016/s0896-6273(00)00138-0
- Ross, L. A., Del Bene, V. A., Molholm, S., Frey, H. P., & Foxe, J. J. (2015). Sex differences in multisensory speech processing in both typically developing children and those on the autism spectrum. *Front Neurosci*, 9, 185. doi:10.3389/fnins.2015.00185
- Rubia, K., Smith, A. B., Taylor, E., & Brammer, M. (2007). Linear age-correlated functional development of right inferior fronto-striato-cerebellar networks during response inhibition and anterior cingulate during error-related processes. *Human Brain Mapping*, 28(11), 1163-1177. doi:10.1002/hbm.20347
- Russchen, F. T., & Jonker, A. J. (1988). Afferent connections of the striatum and the nucleus accumbens in the lizard *Gekko gekko*. *Journal of Comparative Neurology*, 36(1), 49-58. doi:10.1159/000316078
- Schlaggar, B. L., Brown, T. T., Lugar, H. M., Visscher, K. M., Miezin, F. M., & Petersen, S. E. (2002). Functional neuroanatomical differences between adults and school-age children in the processing of single words. *Science*, 296(5572), 1476-1479. doi:10.1126/science.1069464
- Schwertman, N. C., Owens, M. A., & Adnan, R. (2004). A simple more general boxplot method for identifying outliers. *Computational Statistics & Data Analysis*, 47(1), 165-174. doi:10.1016/j.csda.2003.10.012
- Siegel, J. S., Mitra, A., Laumann, T. O., Seitzman, B. A., Raichle, M., Corbetta, M., & Snyder, A. Z. (2017). Data Quality Influences Observed Links Between Functional Connectivity and Behavior. *Cerebral Cortex*, 27(9), 4492-4502. doi:10.1093/cercor/bhw253
- Simonet, M., Roten, F. C. v., Spierer, L., & Barral, J. (2019). Executive control training does not generalize, even when associated with plastic changes in domain-general prefrontal areas. *NeuroImage*, 197, 457-469. doi:10.1016/j.neuroimage.2019.04.010
- Specht, K., Holtel, C., Zahn, R., Herzog, H., Krause, B. J., Mottaghy, F. M., Radermacher, I., Schmidt, D., Tellmann, L., Weis, S., Willmes, K., & Huber, W. (2003). Lexical decision of nonwords and pseudowords in humans: a positron emission tomography study. *Neuroscience Letters*, 345(3), 177-181. doi:10.1016/s0304-3940(03)00494-4
- Sugihara, T., Diltz, M. D., Averbek, B. B., & Romanski, L. M. (2006). Integration of Auditory and Visual Communication Information in the Primate Ventrolateral Prefrontal Cortex. *Journal of Neuroscience*, 26(43), 11138-11147. doi:10.1523/jneurosci.3550-06.2006
- Talsma, D., Senkowski, D., Soto-Faraco, S., & Woldorff, M. G. (2010). The multifaceted interplay between attention and multisensory integration. *Trends in Cognitive Sciences*, 14(9), 400-410. doi:10.1016/j.tics.2010.06.008
- van Atteveldt, N. M., Blau, V. C., Blomert, L., & Goebel, R. (2010). fMR-adaptation indicates selectivity to audiovisual content congruency in distributed clusters in human superior temporal cortex. *BMC Neuroscience*, 11(1), 11-10. doi:10.1186/1471-2202-11-11
- van Atteveldt, N. M., Formisano, E., Blomert, L., & Goebel, R. (2006). The Effect of Temporal Asynchrony on the Multisensory Integration of Letters and Speech Sounds. *Cerebral Cortex*,

- 17(4), 962-974. doi:10.1093/cercor/bhl007
- van Atteveldt, N. M., Formisano, E., Goebel, R., & Blomert, L. (2004). Integration of letters and speech sounds in the human brain. *Neuron*, 43(2), 271-282. doi:10.1016/j.neuron.2004.06.025
- Vannest, J., Karunanayaka, P. R., Schmithorst, V. J., Szaflarski, J. P., & Holland, S. K. (2009). Language Networks in Children: Evidence from Functional MRI Studies. *American Journal of Roentgenology*, 192(5), 1190-1196. doi:10.2214/ajr.08.2246
- Vatansever, D., Menon, D. K., Manktelow, A. E., Sahakian, B. J., & Stamatakis, E. A. (2015). Default Mode Dynamics for Global Functional Integration. *J Neurosci*, 35(46), 15254-15262.
- Visser, M., & Lambon Ralph, M. A. (2011). Differential contributions of bilateral ventral anterior temporal lobe and left anterior superior temporal gyrus to semantic processes. *J Cogn Neurosci*, 23(10), 3121-3131. doi:10.1162/jocn_a_00007
- Wagner, G., de la Cruz, F., Kohler, S., Pereira, F., Richard-Devantoy, S., Turecki, G., Bar, K. J., & Jollant, F. (2019). Connectomics-Based Functional Network Alterations in both Depressed Patients with Suicidal Behavior and Healthy Relatives of Suicide Victims. *Sci Rep*, 9(1), 14330.
- Wang, J., Wang, X., Xia, M., Liao, X., Evans, A., & He, Y. (2015). GRETNA: a graph theoretical network analysis toolbox for imaging connectomics. *Front Hum Neurosci*, 9, 386.
- Wang, X. L., & Tao, B. P. (1996). Chinese character recognition test battery and assessment scale for primary school children. Shanghai Education Press, Shanghai (in Chinese).
- Whitfield-Gabrieli, S., & Nieto-Castanon, A. (2012). Conn: a functional connectivity toolbox for correlated and anticorrelated brain networks. *Brain Connect*, 2(3), 125-141.
- Xia, M., Wang, J., & He, Y. (2013). BrainNet Viewer: a network visualization tool for human brain connectomics. *PLoS One*, 8(7), e68910.
- Xu, W., Kolozsvári, O. B., Oostenveld, R., Leppänen, P. H. T., & Hämäläinen, J. A. (2019). Audiovisual Processing of Chinese Characters Elicits Suppression and Congruency Effects in MEG. *Frontiers in Human Neuroscience*, 13, 18. doi:10.3389/fnhum.2019.00018
- Xu, Y., Pollatsek, A., & Potter, M. C. (1999). The activation of phonology during silent Chinese word reading. *Journal of Experimental Psychology-Learning Memory and Cognition*, 25(4), 838-857. doi:10.1037//0278-7393.25.4.838
- Yang, Y., Yang, Y. H., Li, J., Xu, M., & Bi, H.-Y. (2020). An audiovisual integration deficit underlies reading failure in nontransparent writing systems: An fMRI study of Chinese children with dyslexia. *Journal of Neurolinguistics*, 54, 100884. doi:10.1016/j.jneuroling.2019.100884
- Yates, M., Locker, L., Jr., & Simpson, G. B. (2003). Semantic and phonological influences on the processing of words and pseudohomophones. *Mem Cognit*, 31(6), 856-866. doi:10.3758/bf03196440
- Ye, Z., Rüsseler, J., Gerth, I., & Münte, T. F. (2017). Audiovisual speech integration in the superior temporal region is dysfunctional in dyslexia. *Neuroscience*, 356, 1-10. doi:10.1016/j.neuroscience.2017.05.017
- Yoncheva, Y. N., Zevin, J. D., Maurer, U., & McCandliss, B. D. (2010). Auditory selective attention to speech modulates activity in the visual word form area. *Cereb Cortex*, 20(3), 622-632. doi:10.1093/cercor/bhp129
- Zalesky, A., Cocchi, L., Fornito, A., Murray, M. M., & Bullmore, E. (2012). Connectivity differences in brain networks. *NeuroImage*, 60(2), 1055-1062. doi:10.1016/j.neuroimage.2012.01.068
- Zalesky, A., Fornito, A., & Bullmore, E. T. (2010). Network-based statistic: identifying differences in

brain networks. *NeuroImage*, 53(4), 1197-1207. doi:10.1016/j.neuroimage.2010.06.041

Zeng, L. L., Wang, D., Fox, M. D., Sabuncu, M., Hu, D., Ge, M., Buckner, R. L., & Liu, H. (2014). Neurobiological basis of head motion in brain imaging. *Proceedings of the National Academy of Sciences*, 111(16), 6058-6062. doi:10.1073/pnas.1317424111

Zhang, R., Geng, X., & Lee, T. M. C. (2017). Large-scale functional neural network correlates of response inhibition: an fMRI meta-analysis. *Brain Struct Funct*, 222(9), 3973-3990. doi:10.1007/s00429-017-1443-x

Zhang, S. L., Perfetti, C. A., & Yang, H. (1999). Whole word, frequency-general phonology in semantic processing of Chinese characters. *Journal of Experimental Psychology-Learning Memory and Cognition*, 25(4), 858-875. doi:10.1037/0278-7393.25.4.858